



(21) (A1) 2,107,505  
(22) 1993/10/01  
(43) 1994/04/09

(51) INTL.CL.<sup>5</sup> C12N-015/11; C12N-015/74; C12N-001/21

(19) (CA) **APPLICATION FOR CANADIAN PATENT** (12)

(54) DNA Sequences Conferring Hypertransposability and  
Vectors Constructed Therewith

(72) Baltz, Richard H. - U.S.A. ;  
Solenberg, Patricia J. - U.S.A. ;

(71) Eli Lilly and Company - U.S.A. ;

(30) (US) 07/959,938 1992/10/08

(57) 10 Claims

5,091,605  
90/9'160'S

**BEST AVAILABLE COPY**

Notice: This application is as filed and may therefore contain an  
incomplete specification.



X-9017

ABSTRACT OF THE DISCLOSURE

IR-L5099-10 confers hypertransposing activity on  
5 transposable elements comprising IR-L5099-10. IR-L5099-  
10 is a 57 base pair sequence generated by a spontaneous  
mutation in TN5099-7. Vectors comprising IR-L5099-10  
are useful in a variety of bacteria including but not  
limited to *Streptomyces*, *Chainia*, *Saccharopolyspora*,  
10 *Streptoverticillium* and *Mycobacterium*.

CLAIMS

We claim:

- 5 1. The hypertransposing DNA sequence of Sequence ID No:5 or a functional fragment thereof.
2. The DNA sequence of Claim 1 that is IR-L5099-  
10 10.
3. A transposable element comprising the DNA sequence of Claim 1.
- 15 4. The transposable element of Claim 3 that is a transposon.
5. The transposon of Claim 4 that is Tn5099-10.
- 20 6. A vector comprising the transposon of Claim 5.
7. A host cell containing a transposable element of Claim 3.
- 25 8. The host cell of Claim 7, said host cell being selected from the group consisting of *Streptomyces*, *Streptoverticillium*, *Saccharopolyspora* and *Mycobacterium*.  
i ?
9. A method of introducing a gene of interest into  
30 a host cell, said method comprising transforming said host cell with a vector of Claim 6.

10. The method of Claim 9 wherein said gene of interest is a gene encoding an enzyme in the biosynthetic pathway of a secondary metabolite.

TitleDNA SEQUENCES CONFERRING HYPERTRANSPOSABILITY AND  
VECTORS CONSTRUCTED THEREWITH

5

Transposable elements, such as transposons, insertion sequences, and transposing bacteriophages, are found in both prokaryotic and eukaryotic cells. These elements are capable of insertion events at few or many different sites along the genome. This ability to transpose and insert, even in the absence of DNA homology, has allowed transposable elements to be exploited as powerful genetic tools. Some examples of the utility of such elements include gene mapping, use as promoter probes, construction of insertion mutations, the mediation of intergeneric recombination, the insertion of genetic information into the genome of a recipient strain, and the genetic tagging of specific genes.

20

The development and exploitation of recombinant DNA technology in *Streptomyces* and related genera depends upon the availability of genetic tools for the introduction of foreign DNA or of multiple copies of native DNA into the genome of these cells. This development has been retarded by the very low number of transposable elements presently available for use in *Streptomyces* and related genera such as *Streptoverticillium*, *Saccharopolyspora*, *Chainia* and *Mycobacteria*.

30

The transposable elements available for utilization in *Streptomyces* include: IS110 (Bruton, C. and K. Chater, 1987, Nucleic Acids Res., 15:7053-7065); IS281

- (Sladkova, A. 1986 Mol. Biol., 20:878-881); Tn4556 (Chung, S. 1987 J. Bacteriol., 169:4436-4441); the 2.6 kilobase minicircle from *Streptomyces coelicolor*
- 5 (Lydiate, D., et al., 1986, Mol. Gen. Genet., 203:79-88) and IS493 and transposons derived therefrom.

- Insertion sequence IS493 was isolated from *Streptomyces lividans* using a transposon trap.
- Solenberg, P. and S. Burgett, 1989, J. Bacteriol., 171: 4807-4813. A variety of transposons have been developed utilizing IS493. See Solenberg, P. and S. Burgett, 1989, J. Bacteriol., 171:4807-4813; McHenney, M. and R. Baltz, 1991, J. Bacteriol., 173: 5578-5581; Hahn, D. et al, 1991, J. Bacteriol., 173:5573-5577; and Solenberg, P.
- 10 and R. Baltz 1991 J. Bacteriol., 173:1096-1104. The aforementioned transposons have proven to be useful in genetic manipulations of the various *Actinomycetales*. Transposons derived from IS493 exhibit transposition frequencies of about  $10^{-6}$ .
- 15 Transposon Tn5099, a xylE promoter probe transposon, which was described in Hahn et al., *supra*, was modified to comprise a promoter driving constitutive production of xylE. Transposon Tn5099-7 is the designation for the Tn5099 derivative containing the
- 20 constitutively expressed xylE gene. The constitutive expression of xylE resulted in a selective pressure against xylE production and ultimately led to a deletion in the region of the xylE promoter and coding sequence. The deletion included a portion of the 5' inverted
- 25 repeat of Tn5099-7 and generated a modified inverted repeat with changes in the 50-57 base area (reading left to right from the sense strand). The altered 5' inverted repeat is designated IR-L5099-10 and confers
- 30

hypertransposability on transposons and other elements comprising IR-L5099-10.

- 5 IR-L5099-10, an inverted repeat generated by a spontaneous deletion in Tn5099-7, confers hypertransposing activity on transposable elements constructed therewith.

Brief Description of the Figures

10

For the purposes of this disclosure, the Figures are not drawn exactly to scale.

Figure 1 -- the restriction site and function map of plasmid pCZA141.

- 15 Figure 2 -- the restriction site and function maps of Tn5096 and Tn5097.

Figure 3-- the restriction site and function map of Tn5099-10.

- 20 For purposes of the present invention, as disclosed and claimed herein, the following terms are defined below.

- Am<sup>R</sup> - the apramycin-resistant phenotype or gene conferring same. The Am<sup>R</sup> gene is also designated  
25 aac(3)IV.

Ap<sup>R</sup> - the ampicillin-resistant phenotype or gene conferring same.

- Gene - a DNA segment which encodes a promoter function, a ribosome binding site and a translatable  
30 gene product.

Hm<sup>R</sup> - the hygromycin-resistant phenotype or gene conferring same.

Hypertransposition - transposition by a transposable element at a frequency of  $10^{-5}$  or greater.

Insertion sequence - small (generally less than  
5 2000 bp) elements which encode no known genes unrelated to insertion function.

Inverted Repeats - DNA sequences at the termini of insertion sequences and transposons which facilitate insertion of the insertion sequence or transposon into  
10 host cell DNA. Inverted repeats are abbreviated IR and are sometimes designated IR-L and IR-R to denote which termini of the insertion sequence or transposon they define with IRL being the left (5') repeat and IR-R being the right (3') repeat.

15 Ori - an origin of replication.

Origin of replication - a DNA sequence which allows the replication of an extrachromosomal element in a specific host cell.

Recombinant DNA cloning vector - any autonomously  
20 replicating agent, including, but not limited to, plasmids and phages, comprising a DNA molecule to which one or more additional DNA segments can be or have been added.

Selectable marker - a gene or combination of genes  
25 of known function, the expression of which can be easily determined.

Transformation - the introduction of DNA into a recipient host cell.

Transposable element - a transposon, an insertion  
30 sequence, or a discrete piece of double stranded DNA which is capable of insertion at few or many sites in a genome.



Transposon - a transposable element, generally larger than 2 kb that may contain additional determinants unrelated to insertion function such as antibiotic or heavy metal resistance-conferring genes or biosynthetic genes.

The present invention provides a novel sequence conferring hypertransposing properties and is designated IR-L5099-10. Tn5099-10 comprises IR-L5099-10. Tn5099-10 is illustrative of the transposons which can be constructed using the IR-L5099-10 hypertransposing sequence of the present invention. Tn5099-10 and other vectors comprising the hypertransposing sequence of the present invention are useful for insertionally inactivating chromosomal genes as well as for inserting any sort of foreign DNA into the Streptomyces genome, gene-mapping, promoter probe construction, construction of insertion mutations, and the like. Another aspect of this invention is the use of this transposable element to create other novel transposons. The invention therefore helps to fulfill the long-felt need for such discrete, clearly defined elements for use in Streptomyces, thereby releasing the skilled artisan from reliance upon extra-chromosomal elements or homologous integration. Surprisingly, the IR-L5099-10 element results in transposition frequencies approximately  $10^3$  fold higher than other transposons derived from IS493, including Tn5096 and Tn5099, and thus provides a means for facile achievement of higher transposition frequencies of genes of interest into host organisms.

The steady state transposition frequency of Tn5099-10 was determined as described in Youngman, Philip, Plasmids: A Practical Approach (1987) K. Hardy, Ed., IRL

Press, Inc., Oxford, U.K. at pages 79-103 with modifications for growth in *Streptomyces*. Three replicate experiments using a temperature sensitive delivery vector were performed. The hypertransposing activity of Tn5099-10 was demonstrated by comparison to Tn5096 and Tn5099 under equivalent experimental conditions. The results are summarized in Table 1.

10

Table 1

	Transposon	Transposition Frequency
15	<u>Tn5096</u>	$4.3 \times 10^{-6}$
	<u>Tn5099</u>	less than $10^{-6}$
	<u>Tn5099-10</u>	$2.5 \times 10^{-3}$

20 The experiments summarized above demonstrate that IR-L5099-10 confers a transposition frequency approximately one thousand times greater than conventional IS423 based vectors. Skilled artisans will appreciate that the thousand fold higher transposition frequency affords a facile means for increasing gene dosage (the number of copies of a given gene) in a cell and as such would be of immense benefit in improving production yields of metabolites such as antibiotics or immunosuppressives by increasing the gene dosage of the rate limiting enzyme in the biosynthetic pathway of the metabolite.

25 Tn5099-10 is similar to Tn5099, which has been described in the scientific literature. See Hahn, R. et al., 1991, J. Bact., 173:5573-5577. Tn5099 was

X-9017

-7-

constructed by adding a xylE coding sequence, a  
hygromycin resistance gene and multiple cloning sites to  
IS493. Tn5099 also contains a phage lambda pL promoter  
5 positioned 5' to the hygromycin resistance gene sequence  
to effect transcription of the hygromycin resistance  
gene product in *E. coli*, thereby enabling facile  
identification of *E. coli* transformants. The  
nucleotide sequence of IS493 is taught in Solenberg, P.  
10 and Burgett, S. J., 1991, Bact. 171:4807-4813.

A restriction site and function map of Tn5099-10 is  
provided in Figure 3. The inverted repeats located at  
the termini of Tn5099-10 are in opposite orientation as  
indicated by the arrows of Figure 3. A lambda pL  
15 promoter is positioned 5' to the hygromycin resistance  
gene. The hygromycin resistance gene comprises its  
homologous promoter, which is operable in *Actinomyces*.  
The phage lambda promoter was added for selection of *E.*  
*coli* transformants. The two open reading frames (OrfA  
20 and OrfB) are described in Solenberg and Burgett, *supra*,  
The translation product of OrfB is thought to be a  
transposase. The function of the OrfA gene product  
remains unelucidated. The OrfA functionality is not  
essential for transposition of IS493 based transposons.

25 The nucleotide sequence of Tn5099-10 is provided  
below as Sequence ID No:1.

2107505

X-9017

-8-

GAGCGTTTTT CAACCTCAGG CTGACGGGGT CAGGGGTCAG AGGCGAGGTA  
ACCATCTGCG GTGATAAATT ATCTCTGGCG GTGTTGACAT AAATACCACT  
GGCGGTGATA CTGAGCACAT CAGATCGGTC GTCGAGGTCC ACAGGCCCTC  
5 GGCACGTTTC CGCCGGGCCT CGGCGACCGC CGCGTCGAGG CGCCGGTCGA  
AGGGGCAGTC CTCCACGGGC AGCTCGTGGA GGGCGCGGGC CAGCTCCGCC  
ATCGCCTCGA CCACGGCGAA CCGCTGGTGC TCGGGCCACT CCTCGGCCGC  
CGCGACGCCG GGGACGGCCT CCGTGACGAG CCACGCGGCG GTGTCTCGG  
CACCGCGCTC GACGACGCGG GGGACGGGGA TCGGCGGGGC CTGGCGGCGC  
10 CTCGCCGTCG CAGAACCAGG CGGTGGCGTA CACCGTCGCC TCGGTCGGCC  
CGTAGAGATT GGCATCCCG ACCGCAGCAC CACCGAGAAC GTCCCCGACG  
TGGCCGACCA GCCCGTCATC GTCAACGCCT GACCGCGGTG CGGACAGGCC  
GTGTGCGGAC CGGCCGTGCG GAATTAAGCC GGCCCGTACC CTGTGAATAG  
AGGTCCGCT GTG ACA CAA GAA TCC CTG TTA CTT CTC GAC CGT  
15 ATT GAT TCG GAT GAT TCC TAC GCG AGC CTG CGG AAC GAC CAG  
GAA TTC TGG GAG CCG CTG GCC CGC CGA GCC CTG GAG GAG CTC  
GGG CTG CCG GTG CCG CCG GTG CTG CGG GTG CCC GGC GAG AGC  
ACC AAC CCC GTA CTG GTC GGC GAG CCC GAC CCG GTG ATC AAG  
CTG TTC GGC GAG CAC TGG TGC GGT CCG GAG AGC CTC GCG TCG  
20 GAG TCG GAG GCG TAC GCG GTC CTG GCG GAC GCC CCG GTG CCG  
GTG CCC CGC CTC CTC GGC CGC GGC GAG CTG CGG CCC GGC ACC  
GGA GCC TGG CCG TGG CCC TAC CTG GTG ATG AGC CGG ATG ACC  
GGC ACC ACC TGG CGG TCC GCG ATG GAC GGC ACG ACC GAC CGG  
AAC GCG CTG CTC GCC CTG GCC CGC GAA CTC GGC CGG GTG CTC  
25 GGC CGG CTG CAC AGG GTG CCG CTG ACC GGC AAC ACC GTG CTC

2107505

X-9017

-9-

ACC CCC CAT TCC GAG GTC TTC CCG GAA CTG CTG CGG GAA CGC  
CGC GCG GCG ACC GTC GAG GAC CAC CGC GGG TGG GGC TAC CTC  
TCG CCC CGG CTG CTG GAC CGC CTG GAG GAC TGG CTG CCG GAC  
5 GTG GAC ACG CTG CTG GCC GGC CGC GAA CCC CGG TTC GGC CAC  
GGC GAC CTG CAC GGG ACC AAC ATC TTC GTG GAC CTG GCC GCG  
ACC GAG GTC ACC GGG ATC GTC GAC TTC ACC GAC GTC TAT GCG  
GGA GAC TCC CGC TAC AGC CTG GTG CAA CTG CAT CTC AAC GCC  
TTC CGG GGC GAC CGC GAG ATC CTG GCC GCG CTG CTC GAC GGG  
10 GCG CAG TGG AAG CGG ACC GAG GAC TTC GCC CGC GAA CTG CTC  
GCC TTC ACC TTC CTG CAC GAC TTC GAG GTG TTC GAG GAG ACC  
CCG CTG GAT CTC TCC GGC TTC ACC GAT CCG GAG GAA CTG GCG  
CAG TTC CTC TGG GGG CCG CCG GAC ACC GCC CCC GGC GCC TGA  
CGC CCC GGG TGA AGA GCC CCGTCGGGCG GTGCCTGACG GGGCTTCTCA  
15 GTTAACCTAG GATCCATGGC GGCCGCAATA TTAATCAGGC TGACGGGGTC  
AGGGGTCAGA GGCGGGGTCT GGATCAGCGC GAGCGGTCGT GAGGTAGTAC  
CGGTTGCGCT CGAAGCTCAC TGCCGTCTAG GACAGAAGCT GACCCTAGCT  
CACGCCTACC GTCGCTGCTG GCAAGCAACA TCAGCGAAAG GACCGCG ATG  
CCC CAG CAG ACC AGC CCC GAG CCG GAG GGC TAC ACC ACC GTT  
20 GCG CCC TGG GTC GTC ACC GAC GAC ACG GGG GCT TTC CTC GAC  
TTC GTC ACC CAG GCA TTC GGG GGC GAG GAA CTC GGG CGG GTG  
TCG ACC GAG GAG GGC TTG ATC GGT CAC GCT GAG ATC CGA GTC  
GGC GAC ACC GTC GTG TTG GCC TTC GAC CGA CAC ACG GAC TGG  
CCC GCC ATG CCG AGC CTG CTG CGC GTA TTT GTC GCC GAT GCG  
25 GAC GAG GCG TTT TCC CAG GCC GTC GCA GCT GGC GGC TAT GTC

2107505

X-9017

-10-

GTC ACT TCC CTG GCC GAC GAC GCC TTC GGA CAG CGC GGA GGG  
CGC ATC AAG GAC CCC TTC GGC AAT ATC TGG TGG GTG GTC AGC  
TGC GTT GAG GAC GTC TCT GAG GGC GAG ATG TGG AAG CGG CTG  
5 CAG GAC CCC GTA TAC GCC GAG GCC ATG CGC GTG GCT CAG GAG  
ACA CTC GAC GCC GAG CTC AGC GGT CGA CGC CAC GGG CGC AGC  
AGC GCA CCC GTC AGG ACA ACC AGC TAG ACCCACGACG  
TCGGGGTTCG CATGCGCCAG ACCGGTAGCC AAGGAGCGTC TTCATTCTGC  
TTCGGCAGCT TCGGGGTTGG CCTGACACAG GGGTGAAGCC CCTGGTAGAT  
10 GGGTTTTTCGA CCAAGAAATC CGTCTCTACT AGAGGCTTCAC GTG CTT GTT  
TAC CCC TCC GGG CTG GAC GTG TCC AGT TCT GCC CTC CGC TTC  
CTC TCC GCC CGT CTG CGG GAA CAC CGC CGT GGT CTC GGC ACC  
CGG TGG CGG CGC CTG AGC GCA GGT CGG CAG GCT CTG CTC GCC  
CTC GCC CAT CTC CGC AAC GGT CAT CCG TAT GTG CAG CTC GCG  
15 GCC GGG TTC GGC GTG GGC ACC ACC ACT GCC TAC CGC TAC GTC  
ACC GAG GCA GCT GAG GTT CTG GCC GCC CTT GCC CCG ACG CTC  
GCC GAG GCG GTC CGG GCG GCG TCG ATG AAG GCG TTC GTC CTG  
CTG GAC GGG ACG CTG CTG CCG ATC GAC CGA ATC GCC GCC GAC  
CGG CCT TTC TAC TCG GGC AAG CAC AAG AAA CAC GGG ATG AAC  
20 GTACAG GTG ATC GCA GAT CCT TCC GGC CGC CTG CTG TGG GCC  
TCG CCG GCC CTG CCC GGA GCC GTC CAC GAC GTC CGT GCG GCC  
CGC GAG CAC GGC ATC ATC GAC ACG CTT GCC ACA GCC GAC GTC  
AAC TGC TGG GCA GAC AAG GGC TAC CAG GGC GCC GGA GGC ACG  
GTC CGT GTT CCC TAC CGC GGC CGG TGG GAG ACG CTC TCT GCA  
25 GGG CAG CAG GCC GTC AAT CGA TCC CAC GCG AAG ATC CGG GCG

CTC GTC GAG CAA GCC GTC GCG ACC CTC AAG TCC TGG CGA CTC  
 CTT CGC AGG CTC CGT TGC TCG ACC ACC CGC GTC ACG GTC CTC  
 GTC CAG GCC GTC CTC TGC CTG CAT CAG GCC AGC TCA GCT TGA  
 5 GGATGGAAAACGCTC

The nucleotide sequence of Sequence ID No:1 should be read in context with the restriction site and function map of Tn5099-10, which is provided in Figure 3.

10 The hypertransposing activity of Tn5099-10 has been localized to the first 57 nucleotides of the 5' insertion sequence. These first 57 nucleotides of Sequence ID No:1 are designated IR-L5099-10. The region of Sequence ID No:1 which confers hypertransposing  
 15 activity (IR-L5099-10) includes the sequence identified as IS493-10 splice site in Figure 3, which was generated when the xylE gene was spontaneously deleted from Tn5099-7 to generate Tn5099-10. The sequence of IR-L5099-10 is set forth below as Sequence ID No:5.

20  
 GAGCGTTTTT CAACCTCAGG CTGACGGGGT CAGGGGTCAG AGGCGAGGTA  
 ACCATCT

The combined teachings of Sequence ID No:1,  
 25 Sequence ID No:5 and Figure 3 enable the construction of unlimited varieties of vectors which utilize the hypertransposing function of IR-L5099-10 (Sequence ID No:5). Figure 3 provides restriction endonuclease sites which allow the facile construction of IR-L5099-10  
 30 vectors from IS493 vectors by merely substituting the IR-L5099-10 sequence for the inverted repeat at the left

end (as viewed in Figure 3) of IS493. Sequence ID No:1 provides the nucleotide sequence of Tn5099-10, which comprises from left to right as viewed in Figure 3: the  
5 Inverted Repeat (IR-L5099-10) which is Sequence ID No:5; phage lambda's leftward promoter; a hygromycin resistance gene; a multiple cloning site; OrfA; OrfB and the pristine IR-R of IS493. The restriction site and function map of Figure 3 provides the oligonucleotide  
10 numbers (reading from left to right) which correspond to the restriction endonuclease sites and thereby renders utilization of the components of Sequence ID No:1 to generate vectors comprising Sequence ID No:5 a routine laboratory exercise. Example 1B teaches the synthesis  
15 of a linker which is readily inserted into the unique SauI site of plasmid pCZA141 (NRRL B18417) to generate pCZA141-10. In the event skilled artisans seek additional information on IS493 insertion sequences and transposons derived therefrom reference is directed to  
20 the teachings of Hahn *et al.*, *supra*; Solenberg and Baltz, *supra*; McHenney and Baltz, *supra*; and Solenberg and Burgett, *supra*, the entire teachings of which are herein incorporated by reference.

Example 1 teaches construction of transposons  
25 derived from the IS493 sequence. Vectors which transpose at higher frequency are readily prepared merely by substituting the IR-L5099-10 sequence of Sequence ID 5 for the corresponding region in vectors comprising IS493. References to literature describing  
30 Tn5096 and Tn5092 have been provided in an earlier section. Skilled artisans will appreciate the facile nature of converting such related transposons into IR-L5099-10 vectors.



The IS493 transposable element can be obtained from plasmid pZCA141, which can be conventionally isolated from *E. coli* K12 JM109/pCZA141, a strain deposited and made part of the permanent stock culture collection of the NRRL. A culture of *E. coli* K12 JM109/pCZA141 can be obtained from the NRRL under the accession number NRRL B18417. A restriction site and function map of plasmid pCZA141 is presented in Figure 1 of the accompanying drawings.

The IS493 element comprises two distinct open reading frames. The first open reading frame, designated OrfA, encodes 177 amino acids and is found between the unique SauI site and the unique SphI site. The second open reading frame, designated OrfB, encodes 233 amino acids and is located at the opposite end of the IS493 element from OrfA. The apramycin resistance-conferring gene (aac(3)IV) ligated into this site can be oriented in either direction. Tn5096, found in plasmid pCZA156, contains the aac(3)IV gene in an orientation opposite of OrfA and OrfB. Tn5097, found in plasmid pCZA157, contains the aac(3)IV gene in the same orientation as OrfA and OrfB. Both Tn5096 and Tn5097 are useful in integrating this foreign gene into *Actinomycetales* genomes. Many other novel transposons can be created by ligating any foreign gene of interest into IS493 or IR-L5099-10 based vectors. Restriction site and function maps of Tn5096 and Tn5097 are found in Figure 2 of the accompanying drawings. Transposition events, such as those made possible by IS493 or the corresponding IR-L5099-10 vectors, are very useful because transposition may occur in a variety of locations in the genome and does not require DNA

homology for integration. In addition, integration of DNA by transposition mechanisms is not limited to organisms that have functional homologous recombination systems.

The IR-L5099-10 hypertransposing sequence is in no way limited to the use in the plasmids herein described, as many different techniques and cloning vectors could be used to introduce the transposable element into the host cell. Furthermore, the types of genes integrated into the chromosome via transposition using the IR-L5099-10 element are also virtually limitless. The antibiotic resistance-conferring genes of the exemplified plasmids could be replaced by many different antibiotic resistance conferring genes or the antibiotic resistance-conferring genes could be replaced by heavy metal resistance-conferring genes, secondary metabolite biosynthetic genes, combinations thereof. The stable integration of secondary metabolite biosynthetic genes into the chromosome of a host cell can cause increased production of the desired gene product or the production of new, hybrid gene products. Those skilled in the art recognize that virtually any gene encoding any function can be transposed into the host chromosome using the IR-L5099-10 element disclosed herein. Therefore, the addition of any extraneous DNA within the parameters of the IR-L5099-10 element would be within the scope of the present invention.

Many modifications and variations of the present illustrative DNA sequence are possible. Such modifications can introduce new restriction sites into the IR-L5099-10 vectors and could greatly increase the utility of the element. Indeed, some modifications

could potentially alter the frequency of transposition controlled by IS493 derivatives. These modified IR-L5099-10 elements may occur naturally such as

5 IR-L5099-10 which generated Tn5099-10 or can be produced by synthetic methods in substantial accordance with the procedures of Itakura *et al.*, 1977, *Science*, 198:1056 and Crea *et al.*, 1978, *Proc. Nat. Acad. Sci.*, USA 75:5765. In addition, synthetic components or linkers

10 can be synthesized either by using a Systec 1450A DNA synthesizer (Systec Inc., 3816 Chandler Drive, Minneapolis, MN) or an ABS 380A DNA synthesizer (Applied Biosystems, Inc., 850 Lincoln Center Drive, Foster City, CA 94404). Many other DNA synthesizing instruments are

15 known in the art and can be used to make synthetic DNA fragments. Therefore, the present invention is in no way limited to the DNA sequences and plasmids specifically exemplified.

Tn5099-10 resulted from the spontaneous deletion of

20 the xylE coding region and its constitutive promoter from a transposon designated Tn5099-7. The deletion eliminated a sequence in the 50-57 base area. See Figure 3 and Sequence ID No:5. Those of ordinary skill in the art realize that further deletions in this area

25 would lead to IRs which would likewise confer hypertransposing activity to insertion sequences and transposons. The availability of aforementioned and numerous other DNA synthesizers as well as the growing numbers of commercial laboratories which perform custom

30 DNA synthesis and exonucleases which "chew-back" the ends of DNA sequences allows the facile preparation of IR-L5099-10-like sequences which are smaller than the IR-L5099-10 sequence of Sequence ID No:5. Accordingly,

the present invention embraces functional subfragments of Sequence ID No:5.

- Plasmid pCZA141-10 is illustrative of the vectors which can be constructed utilizing the IR-L5099-10 sequence of the present invention. Plasmid pCZA141-10 contains Tn5099-10 which has IR-L5099-10 as the IR-L and the IR-R of IS493. Skilled artisans realize that the IR-R of Tn5099-10 can be readily replaced with IR-L5099-10 to generate further hypertransposing transposons. Skilled artisans also realize that the non-Tn5099-10 components of plasmid pCZA141-10 are readily interchangeable with other vector components which provide the same functions. The non-Tn5099-10 components of plasmid pCZA141-10 are derived from pUC19, a commercially available vector with numerous functional equivalents also being available from depositories such as the ATCC and the NRRL as well as numerous commercial concerns such as New England Biolabs, Pharmacia, Boehringer Mannheim, etc.. Accordingly, the vector backbone (non-Tn5099-10) is not a limitation to the vectors which can be used to introduce the IR-L5099-10 transposable elements of the present invention into host cells.
- The transposable elements constructed with the IRL-5099-10 sequence are useful in numerous species and strains of *Streptomyces*, *Chainia*, *Saccharopolyspora*, *Streptovercillium* and *Mycobacterium*. The transformation and transfection protocols required for introduction of vectors into the aforementioned bacteria are well known in the art and will not be reproduced as part of this disclosure.

The examples which follow are intended to further illustrate the present invention and are not to be interpreted as limiting on the scope thereof. While the examples and detailed description sections of the present invention are sufficient to guide anyone of ordinary skill in the art in the practice of the present invention, skilled artisans are also directed to Molecular Cloning A Laboratory Manual, Second Edition, Sambrook, J., Fritsch, E. F., and Maniatis, T., Cold Spring Harbor Press, 1989, and Current Protocols In Molecular Biology, Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D. D., Seidman, J.G., Smith, J.A., and Struhl, K., Ed. Greene Publishing Associates and Wiley-Interscience, 1989. The aforementioned resources provide an excellent technical supplement to any discourse in genetic engineering. *Streptomyces* and related genera such as *Saccharopolyspora*, *Streptoverticillium*, and *Chainia* require adaptation of the techniques described in Sambrook, *supra* and Molecular Cloning, which are directed primarily towards *E. coli* and related gram negative bacilli. Genetic Manipulation of Streptomyces-A Laboratory Manual, Hopwood, D. A., et al., Ed. F. Crowe and Sons, Ltd., Norwich, U.K. provides an excellent supplement as it provides a "methods book" approach to the molecular biology protocols used in genetic manipulations of *Streptomyces*.

Example 1CONSTRUCTION OF PLASMID pCZA141-10

5

A. Isolation of Plasmid pCZA141

Plasmid pCZA141, which contains IS<sub>493</sub> on an approximately 2.1 kb HindIII restriction fragment is obtained from the Northern Regional Research Center, NRRL, from *E. coli* K12 JM109/pCZA141 under the accession number NRRL B-18417. A restriction site and function map of plasmid pCZA141 is presented in Figure 1 of the accompanying drawings.

The lyophils of K12 JM109/pCZA141 are plated onto L-agar plates (10 g of Bacto-tryptone, 10 g of NaCl, 5 g of Bacto-Yeast Extract, and 15 g of agar per liter) containing 25 µg/ml ampicillin to obtain a single colony isolate of the strain. This colony is used to inoculate about 500 ml of L broth (1% (w/v) Bacto tryptone; 0.5% (w/v) yeast extract; 0.5% NaCl and 0.1% (w/v) dextrose at pH 7) containing 25 µg/ml ampicillin, and the resulting culture is incubated at 37°C with aeration until the cells reach stationary phase.

Plasmid DNA is obtained from the cells in accordance with the following procedure, which is adapted from Maniatis et al., 1982, Molecular Cloning (Cold Spring Harbor Laboratory). This same procedure is used, but on a smaller scale and with the ultracentrifugation steps replaced with phenol followed by chloroform extractions, to prepare the plasmid DNA used to identify the *E. coli* transformants.

About 500 ml of stationary-phase *E. coli* K12 JM109/pCZA141 cells are harvested by centrifugation at

4000Xg for 10 minutes at 4°C, and the supernatant is discarded. The cell pellet is washed in 100 ml of ice-cold STE buffer (0.1 M NaCl; 10 mM Tris-HCl, pH 7.8; and 1 mM EDTA). After the cell pellet is washed, the pellet is resuspended in 10 ml of Solution 1 (50 mM glucose; 25 mM Tris-HCl, pH=8.0; and 10 mM EDTA) that contains 5 mg/ml lysozyme and is left at room temperature for 10 minutes. Twenty ml of Solution 2 (0.2 N NaOH and 1% SDS) are then added to the lysozyme-treated cells, and the solution is gently mixed by inversion. The mixture is incubated on ice for 10 minutes.

Fifteen ml of ice-cold, 3 M sodium acetate, pH=4.8, are added to the lysed-cell mixture, and the solution is mixed by inversion. The solution is incubated on ice for 60 minutes. The 3 M sodium acetate solution is prepared by mixing equal volumes of 3 M acetic acid and 3 M sodium acetate.

The lysed cell mixture is centrifuged in a Beckman SW27 rotor (or its equivalent) at 20,000 rpm for 20 minutes at 4°C. About 36 ml of supernatant are recovered, and 2.5 volumes of ethanol are added, mixed, and the resulting solution left on ice for 15 minutes. The plasmid DNA is collected by centrifugation at 12,000Xg for 30 minutes at room temperature. The supernatant is discarded, and the DNA pellet is washed with 70% ethanol at room temperature. The ethanol wash is decanted, and the pellet is dried in a vacuum desiccator. The pellet is then resuspended in 8 ml of TE buffer.

Eight grams of CsCl are added to the DNA solution. About 0.8 ml of a 10 mg/ml solution of ethidium bromide

in water are added for each 10 ml of CsCl-DNA solution. The final density of the solution is about 1.55 g/ml, and the ethidium bromide concentration is about 800 µg/ml. The solution is transferred to a Beckman Type 50 centrifuge tube, filled to the top with TE buffer containing 1.55 g/ml CsCl, sealed, and centrifuged at 45,000 rpm for 24 hours at 20°C. After centrifugation, two bands of DNA are visible in ordinary light and become even more prominent in UV light. The cap is removed from the tube, and the lower DNA band is recovered using a syringe with a #21 hypodermic needle inserted through the side of the centrifuge tube.

The ethidium bromide is removed from the solution of plasmid DNA by several extractions with water-saturated 1-butanol, and the CsCl is removed by dialysis against TE buffer. After extractions with buffered phenol and then chloroform, the DNA is precipitated, washed with 70% ethanol, and dried. About 0.5 mg of plasmid pCZA141 DNA can be obtained by this procedure.

B. Digestion of Plasmid pCZA141 with SauI

About 5 µg of plasmid pCZA141 are digested with SauI. The composition of digestion buffers and reaction conditions for restriction endonuclease cleavage are provided in New England Biolabs Catalogs and product literature as well as the Sambrook *et al.*, *supra*, and Molecular Cloning, *supra*, references and will not be recited here. Plasmid pCZA141 contains a single SauI site, which is near the 5' end of IS493 and within the IR-L of IS493. Digestion of plasmid pCZA141 linearizes the vector to accommodate the DNA linker which convert



X-9017

-21-

IS493 to Tn5099-10. Following digestion with SauI the linear pCZA141 was conventionally purified.

5 C. Preparation of the Linker

Ligation of the DNA sequence provided below into the unique SauI site of plasmid pCZA141 IS493 generates Tn5099-10.

```
10 TCAGGCTGAC GGGGTCAGGG GTCAGAGGCG AGGTAACCAT CTGCGGTGAT
    AAATTATCTC TGGCGGTGTT GACATAAATA CCACTGGCGG TGATACTGAG
    CACATCAGAT CGGTCGTCGA GGTCCACAGG CCCTCGGCGA CGTTCCGCCG
    GGCCTCGGCG ACCGCCGCGT CGAGGCGCCG GTCGAAGGGG CAGTCCTCCA
    CGGGCAGCTC GTGGAGGGCG CGGGCCAGCT CCGCCATCGC CTCGACCACG
15 GCGAACCCTG GGTGCTCGGG CCACTCCTCG GCCGCCGCGA CGCCGGGGAC
    GGCCTCCGTG ACGAGCCACG CGGCGGTGTC GTCGGCACCG CGCTCGACGA
    CGCGGGGGAC GGGGATCGGC GGGGCCTGGC GGCGCCTCGC CGTCGCAGAA
    CCAGGCGGTG GCGTACACCG TCGCCTCGGT CGGCCCCGTAG AGATTGGCGA
    TCCCGACCGC AGCACCACCG AGAACGTCCC CGACGTGGCC GACCAGCCCC
20 TCATCGTCAA CGCCTGACCG CGGTGCGGAC AGGCCGTGTC GCGACCGGCC
    GTGCGGAATT AAGCCGGCCC GTACCCTGTG AATAGAGGTC CGCT    GTG
    ACA CAA GAA TCC CTG TTA CTT CTC GAC CGT ATT GAT TCG GAT
    GAT TCC TAC GCG AGC CTG CGG AAC GAC CAG GAA TTC TGG GAG
    CCG CTG GCC CGC CGA GCC CTG GAG GAG CTC GGG CTG CCG GTG
25 CCG CCG GTG CTG CGG GTG CCC GGC GAG AGC ACC AAC CCC GTA
    CTG GTC GGC GAG CCC GAC CCG GTG ATC AAG CTG TTC GGC GAG
    CAC TGG TGC GGT CCG GAG AGC CTC GCG TCG GAG TCG GAG GCG
```

TAC GCG GTC CTG GCG GAC GCC CCG GTG CCG GTG CCC CGC CTC  
CTC GGC CGC GGC GAG CTG CGG CCC GGC ACC GGA GCC TGG CCG  
TGG CCC TAC CTG GTG ATG AGC CGG ATG ACC GGC ACC ACC TGG  
5 CGG TCC GCG ATG GAC GGC ACG ACC GAC CGG AAC GCG CTG CTC  
GCC CTG GCC CGC GAA CTC GGC CGG GTG CTC GGC CGG CTG CAC  
AGG GTG CCG CTG ACC GGG AAC ACC GTG CTC ACC CCC CAT TCC  
GAG GTC TTC CCG GAA CTG CTG CGG GAA CGC CGC GCG GCG ACC  
GTC GAG GAC CAC CGC GGG TGG GGC TAC CTC TCG CCC CGG CTG  
10 CTG GAC CGC CTG GAG GAC TGG CTG CCG GAC GTG GAC ACG CTG  
CTG GCC GGC CGC GAA CCC CGG TTC GGC CAC GGC GAC CTG CAC  
GGG ACC AAC ATC TTC GTG GAC CTG GCC GCG ACC GAG GTC ACC  
GGG ATC GTC GAC TTC ACC GAC GTC TAT GCG GGA GAC TCC CGC  
TAC AGC CTG GTG CAA CTG CAT CTC AAC GCC TTC CGG GGC GAC  
15 CGC GAG ATC CTG GCC GCG CTG CTC GAC GGG GCG CAG TGG AAG  
CGG ACC GAG GAC TTC GCC CGC GAA CTG CTC GCC TTC ACC TTC  
CTG CAC GAC TTC GAG GTG TTC GAG GAG ACC CCG CTG GAT CTC  
TCC GGC TTC ACC GAT CCG GAG GAA CTG GCG CAG TTC CTC TGG  
GGG CCG CCG GAC ACC GCC CCC GGC GCC TGACGCCCCG  
20 GGTGAAGAGC CCCGTCGGGC GGTGCCTGAC GGGGCTTCTC AGTTAACCTA  
GGATCCATGG CGGCCGCAAT ATTAA (Sequence ID No:6)

Oligonucleotide synthesis is well known in the art and the sequence of IR-L5099-10 (Sequence ID No:6) is readily synthesized by such methods. See Sambrook and Maniatis, supra, and Current Protocols in Molecular Biology, supra. The linker sequence corresponds to Sequence ID No:6. The linker is synthesized as a series of oligonucleotides which collectively encompass both strands of Sequence ID No:6. The selection of appropriate oligonucleotides is a matter of choice. The present inventors did not synthesize the fragment since Tn5099-10 resulted from a spontaneous deletion in Tn5099-7 as described in an earlier section of the specification. The linker provided herein is provided to supplement the combined teachings of Sequence ID No:1 (IS493) and Sequence ID No:5 (IR-L5099-10), which are deemed enabling in and of themselves.

D. Ligation of the 1675 bp linker into pCZA141 to generate plasmid pCZA141-10

The linearized pCZA141 of Example 1A, which has SauI overhangs and the 1675 bp linker of Example 1B, which also has SauI overhangs are ligated in accordance with procedures well known in the art. See Sambrook et al., *supra*, at 1.53-1.72.

E. Transformation of *E. coli* MM294 with the ligation mixture of Example 1C.

To prepare *E. coli* K12 MM294 cells that are competent for transformation, the lyophils of *E. coli* K12 MM294 obtained from the NRRL (accession number NRRL B-15625) are reconstituted to isolate single colonies. One single-colony isolate of MM294 is inoculated into

5 ml of broth and the culture is incubated at 37°C overnight with aeration. Fifty  $\mu$ l of the overnight culture are used to inoculate 5 ml of broth. The culture is incubated at 37°C overnight with aeration. The following morning, the culture is diluted to 200 ml with broth. The diluted culture is incubated at 37°C with aeration until the absorbance at 550 nm (A) is about 0.5. The culture is cooled for ten minutes in an ice-water bath, and the cells are then collected by centrifugation at 4000Xg for 10 minutes at 4°C. The cell pellet is resuspended in 100 ml of cold 10 mM NaCl and then immediately repelleted by centrifugation. The cell pellet is resuspended in 100 ml of 30 mM CaCl and incubated on ice for 20 minutes.

The cells are again collected by centrifugation and resuspended in 10 ml of 30 mM CaCl. A one-half ml aliquot of the cells is added to the ligated DNA prepared above. The cell-DNA mixture is incubated on ice for one hour, heat-shocked at 42°C for 90 seconds, and then chilled on ice for about two minutes. The cell-DNA mixture is centrifuged, and the cell pellet is resuspended into 0.5 ml of broth in a 1.5 ml tube and incubated at 37°C for one hour.

Aliquots of the transformation mixture are plated on L-agar plates containing 100  $\mu$ g apramycin/ml. The plates are incubated at 37°C overnight. Several apramycin-resistant colonies are selected and then screened by restriction enzyme analysis. A restriction site and function map of plasmid pCZA141 is provided in Figure 1. The insertion of the 1675 bp linker of Example 1B into the unique SauI site results in two different plasmids which differ only in the orientation

X-9017

-25-

of the linker. The plasmid having the linker in proper orientation contains Tn5099-10. Digestion of the desired plasmid with EcoRI and SauI results in DNA

5 fragments of 3.6 kb and 2.9kb when resolved on an agarose gel. The plasmid having the linker in proper orientation and thus comprising Tn5099-10 (Sequence ID No:1) is designated pCZA141-10. Restriction endonuclease mapping is discussed in Sambrook et al., supra, at 1.85

10 et. seq.

FIG. 1

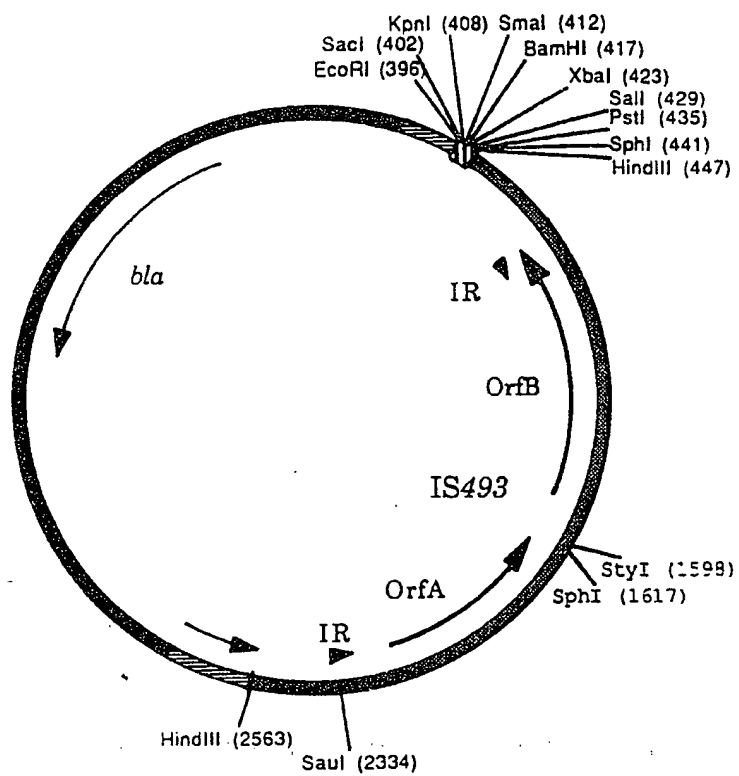


FIG. 2

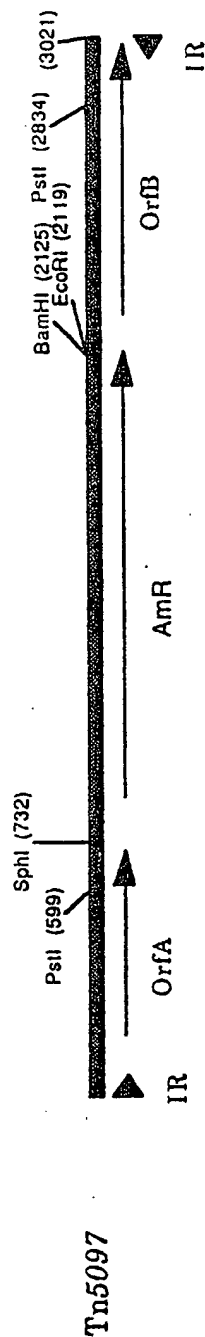
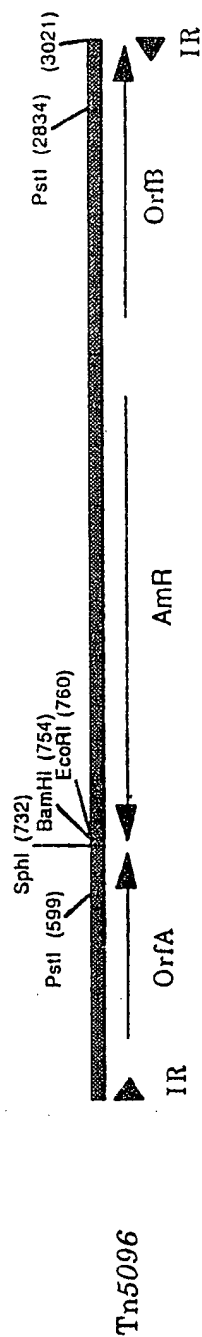
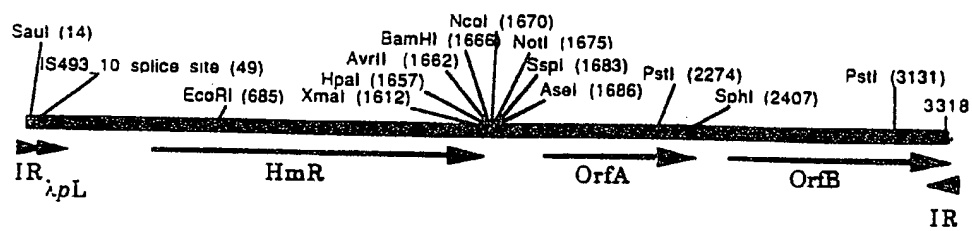


FIG. 3



IS493\_10  
hypertransposon



**This Page is Inserted by IFW Indexing and Scanning  
Operations and is not part of the Official Record**

**BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☒ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: \_\_\_\_\_

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.**